CORTEX 63 (2015) 196–205

Available online at www.sciencedirect.com

ScienceDirect

Journal homepage: www.elsevier.com/locate/cortex

Research report

Sound symbolism scaffolds language development in preverbal infants

Michiko Asano [a,b,c], Mutsumi Imai [b,*], Sotaro Kita [d], Keiichi Kitajo [e,f], Hiroyuki Okada [c] and Guillaume Thierry [g]

[a] Department of Psychology, College of Contemporary Psychology, Rikkyo University, Niiza, Saitama, Japan
[b] Faculty of Environment and Information Studies, Keio University, Fujisawa, Kanagawa, Japan
[c] Tamagawa University Brain Science Institute, Machida, Tokyo, Japan
[d] Department of Psychology, University of Warwick, Coventry, UK
[e] RIKEN BSI-Toyota Collaboration Center, RIKEN Brain Science Institute, Wako, Saitama, Japan
[f] RIKEN Brain Science Institute, Wako, Saitama, Japan
[g] School of Psychology, Bangor University, Gwynedd Wales, UK

ARTICLE INFO

Article history:
Received 1 February 2014
Reviewed 10 April 2014
Revised 22 June 2014
Accepted 26 August 2014
Action editor Anne Castles
Published online 16 September 2014

Keywords:
Sound symbolism
Language development
Audio-visual correspondences
Phase synchronization analysis of EEG
Amplitude change analysis of EEG

ABSTRACT

A fundamental question in language development is how infants start to assign meaning to words. Here, using three Electroencephalogram (EEG)-based measures of brain activity, we establish that preverbal 11-month-old infants are sensitive to the non-arbitrary correspondences between language sounds and concepts, that is, to sound symbolism. In each trial, infant participants were presented with a visual stimulus (e.g., a round shape) followed by a novel spoken word that either sound-symbolically matched ("moma") or mismatched ("kipi") the shape. Amplitude increase in the gamma band showed perceptual integration of visual and auditory stimuli in the match condition within 300 msec of word onset. Furthermore, phase synchronization between electrodes at around 400 msec revealed intensified large-scale, left-hemispheric communication between brain regions in the mismatch condition as compared to the match condition, indicating heightened processing effort when integration was more demanding. Finally, event-related brain potentials showed an increased adult-like N400 response – an index of semantic integration difficulty – in the mismatch as compared to the match condition. Together, these findings suggest that 11-month-old infants spontaneously map auditory language onto visual experience by recruiting a cross-modal perceptual processing system and a nascent semantic network within the first year of life.

* *Corresponding author*. Faculty of Environment and Information Studies, Keio University, 5322 Endo, Fujisawa, Kanagawa 252-0882, Japan.
E-mail address: imai@sfc.keio.ac.jp (M. Imai).
http://dx.doi.org/10.1016/j.cortex.2014.08.025

1. Introduction

How human infants map speech sounds to meaning in order to break into semantics is a key question for understanding the ontogenesis of language. It has been suggested that a biologically endowed ability to realize cross-modal mapping, particularly between auditory and visual percepts, scaffolds language learning in human infants (Imai, Kita, Nagumo, & Okada, 2008; Maurer, Pathman, & Mondloch, 2006). Consistent with this idea, 4-month-old infants appear to sense intrinsic correspondences between speech sounds and certain features of visual input (see Ozturk, Krehm, & Vouloumanos, 2013; Peña, Mehler, & Nespor, 2011), a phenomenon referred to as sound symbolism. It has also been reported that toddlers are not only sensitive to sound symbolism (Maurer et al., 2006) but also make use of sound symbolism in verb learning (Imai et al., 2008; Kantartzis, Imai, & Kita, 2011). The results from preverbal infants (Ozturk et al., 2013; Peña et al., 2011) and those from toddlers (Imai et al., 2008; Kantartzis et al., 2011; Maurer et al., 2006) support the idea that sound symbolism plays an important role in the ontogenesis of language (Imai & Kita, 2014; Imai et al., 2008; Maurer et al., 2006).

It is generally agreed that infants start to associate speech sounds and visual referents at around 12–14 months. At this age, however, the process is effortful because infants have limited information processing capacities and little experience in mapping words to the world (Fennell & Werker, 2003; Werker, Cohen, Lloyd, Casasola, & Stager, 1998). They may rely more on perceptually based cues that are available without prior word learning experiences, such as cross-modal correspondences between speech and visual input in their word learning. Indeed, previous research suggests that 14-month-old infants use sound-symbolic correspondences between speech sounds and object properties as a cue in their effort to establish word (speech sounds) – referent associations (Imai et al., under review; Miyazaki et al., 2013). Thus, there is some evidence that sound symbolism helps young infants at the initial stage of word learning.

However, how sound symbolism is processed in the infants' brain has not yet been addressed in the literature. It is not conceivable that four-month-old infants are actively engaged in semantic processing when they hear speech sounds together with a visually presented referent (Stager & Werker, 1997). Thus, infants at this age are likely to process sound symbolism perceptually, possibly on the basis of cross-modal binding mechanisms. However, at later times, the influence of sound symbolism is likely to transpire in temporal windows compatible with higher-level information processing, i.e., the semantic level.

In this study, we investigated how 11-month-old infants respond to sound symbolism. If perceptual cross-modal mapping ability scaffolds the establishment of word-referent associations, we might see the effect of sound symbolism in two time-windows: (a) in an early time window coinciding with the time period of perceptual processing, and (b) in a time window coinciding with higher-level cognition and/or semantic processing. We chose to study 11-month-olds because they are just about to say their first words but there is little or no evidence to date for the successful establishment of novel word-referent associations in experimental settings at this age. To examine this possibility, we recorded EEG from nineteen children during the presentation of novel word – visual shape pairs that were either sound-symbolically congruent or incongruent (Fig. 1) and analysed the data using three indices of brain functions: amplitude change, large-scale phase synchronization, and event-related potentials (ERPs). In each trial, infants were presented with a picture of a shape (randomly selected from 20 spiky and 20 round shapes) followed by a novel word ("kipi" or "moma").

Here, we were interested in testing whether infants would manifest increased N400 amplitude in the case of sound-symbolically mismatching word-shape pairs as compared to sound-symbolically matched ones. The N400 effect is an ERP modulation known to be sensitive to semantic integration processes in adults (Kutas & Federmeier, 2011), but also in infants (Friedrich & Friederici, 2005, 2011; Parise & Csibra, 2012). A more negative-going N400 deflection for sound symbolically mismatching sound-shape pairs would indicate that infants with very little vocabulary assume sound symbolic correspondence between word sound and shape, and consider sound-shape mismatches to be anomalies at a conceptual/semantic level.

Accumulating evidence suggests that an increase in gamma-band EEG amplitude, or gamma-band activity, is related to cross-modal perceptual integration. For example, Schneider, Debener, Oostenveld, and Engel (2008) reported that gamma-band activity increased for matched audio-visual stimuli at around 100–200 msec in the 40–50 Hz frequency range in adults (see also Senkowski, Schneider, Foxe, & Engel, 2008 for a review).

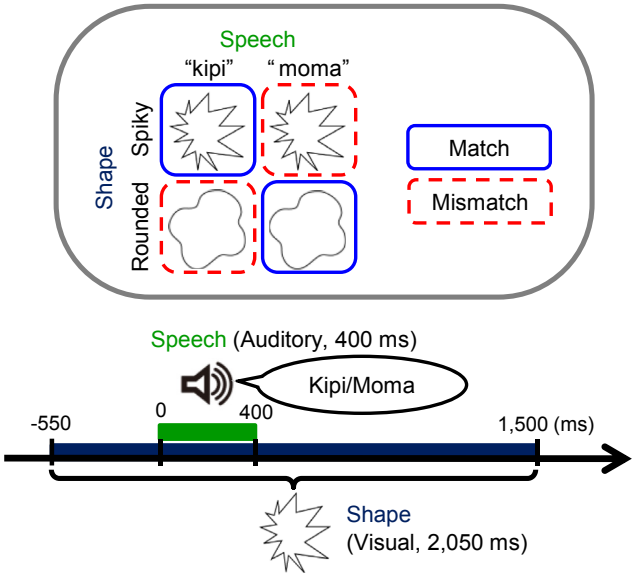

Fig. 1 – Experimental protocol. In each trial, infants were presented with a spiky or round visual shape, followed by a novel word consisting of either voiceless stops and high vowels or nasal consonants and mid/low vowels ("kipi" and "moma", respectively). Shape and sound were sound-symbolically matched (e.g., a spiky shape followed by "kipi") or mismatched (e.g., a spiky shape followed by "moma").

In the present study, we analysed amplitude changes, especially in the gamma-band to investigate whether infants process sound symbolism perceptually within local networks underpinning cross-modal perceptual integration. To our knowledge, no previous study has shown how infants' cross-modal processing is reflected in amplitude changes. However, previous studies have demonstrated that gamma-band activity is related to uni-modal perceptual binding both in adults (cf. Tallon-Baudry, Bertrand, Delpuech, & Pernier, 1996) and infants (cf. 8-month-olds, Csibra, Davis, Spratling, & Johnson, 2000). These results suggest that gamma-band activity might be related to perceptual binding in infants, either within one or across different modalities. Thus, here we may see the gamma-band amplitude changes in a similar time window if sound symbolism is processed as cross-modal binding between audition and vision.

Large-scale synchronization of neural oscillations has been shown to play an important role in the dynamic linking of distributed brain regions in adults (Engel & Singer, 2001; Fries, 2005; Kawasaki, Kitajo, & Yamaguchi, 2010; Kitajo et al., 2007; Lachaux et al., 2000; Rodriguez et al., 1999; Varela, Lachaux, Rodriguez, & Martinerie, 2001; Ward, 2003). Semantic processing requires communication between distributed brain regions; thus, such exchange should be further reflected in large-scale phase synchronization of neural activity. Here, for the first time, we investigated large-scale phase synchronization to study how different areas of the infant brain communicate, when children process sound symbolically matching and mismatching novel word-shape pairs.

We expected that each of the three analyses would index different aspects of sound symbolism and allow us to gain a better and deeper understanding of infants' neural activities relating to meaning integration. We thus focused on how the results from the three analyses could be related and complement one another.

2. Materials and methods

2.1. Participants

Forty-nine healthy Japanese 11-month-old infants participated in this experiment. Informed consent was obtained from all participants (parents of the infants and adults participated in the rating studies) of this study after the nature and possible consequences of the studies were explained, and the rights of the participants were protected. All the experimental procedures had been approved by the Ethical Committee of Tamagawa University, Japan, where the experiment was carried out. We included only those infants who had a minimum of 20 artefact-free trials per condition. Data from 30 infants were excluded from the analyses because of fussiness ($N = 23$) or insufficient data ($N = 7$). A total of 19 infants (13 boys, 6 girls, $M = 11$ months and 25 days, range = 11 months and 6 days to 12 months and 22 days) entered the final analyses.

2.2. Stimuli

Twenty spiky shapes and twenty rounded shapes, drawn with black lines on a white background, were prepared. Stimulus words and shapes were selected on the basis of the literature on shape sound symbolism (Köhler, 1947; Maurer et al., 2006; Ramachandran & Hubbard, 2001) and pretests. Each image was presented to infants four times (twice with the matched sound and twice with the mismatched sound) resulting in 160 randomly ordered trials. In each trial, participants were shown one of the spiky or rounded visual shapes, followed by one of two nonsense words, "kipi" and "moma", spoken by a Japanese female (400 msec in duration). These words and shapes were selected on the basis of the literature on shape sound symbolism (Köhler, 1947; Maurer et al., 2006) and pretests. The degree of sound-symbolic match for each combination of shapes and words was highly ranked in pretests including other word-shape pairs in adult speakers of Arabic ($N = 18$), Japanese ($N = 98$) and English ($N = 83$). Examples of the shapes are shown in Fig. 2.

2.3. Procedure

Infants were seated on the lap of a caregiver and tested in front of a 37 inch liquid crystal display (SHARP AQUOS LC-37DS5 set to a 1280 × 1024 pixels resolution with a 60 Hz refresh rate) in an electrically shielded and sound attenuated room. The viewing distance was about 1.2 m. Caregivers wore headphones to prevent them from hearing the auditory stimuli and potentially influence their child's behaviour. Each trial was initiated manually to insure that the infant's attention was directed towards the screen. A visual shape appeared in the centre of the display within a viewing angle of approximately 7 × 7°, 200 msec after the beginning of the trial and remained for 2050 msec. The spoken word 'kipi' or 'moma' (400 msec in duration) was presented 550 msec after the onset of the visual stimulus. Infants passively saw and heard the stimuli. An attention-getter was presented in one fourth of the trials (randomly selected) to regularly reinforce the infants' attention towards the display.

2.4. Electroencephalogram (EEG) recording

The EEGs were continuously recorded from silver–silver chloride electrodes attached to an elastic electrode cap. EEG data were recorded at 11 electrode sites: F3, Fz, F4, C3, Cz, C4, P3, Pz, P4, and left and right mastoids (A1, A2). The ground electrode was placed at FPz. Electrode impedances were kept mostly below 10 kΩ. The EEG activity was amplified with Neuroscan SynAmps2, digitized online at a rate of 1 kHz, and filtered on-line (bandpass between .1 and 200 Hz).

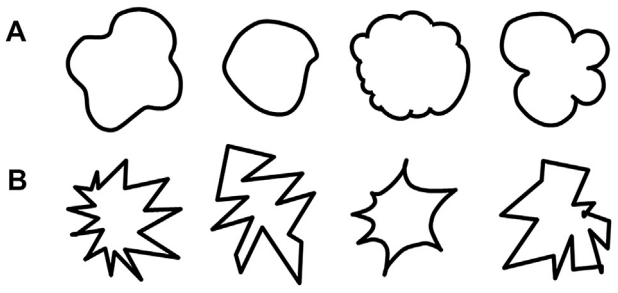

Fig. 2 – Randomly selected examples of the variants of (a) rounded and (b) spiky stimulus shapes.

2.5. Calculation

2.5.1. Amplitude change

The EEG was re-referenced to the average of left and right mastoid channels (A1, A2). Artifact rejection was performed based on the criteria used in the ERP analyses (see section 2.5.2). There was a minimum of 21 valid epochs per condition in every infant participant (mean: 47.6 epochs in the match condition and 46.7 epochs in the mismatch condition). Epochs ranged from −2000 to 1500 msec after the auditory onset.

To estimate local brain networks, we extracted amplitude of oscillations in each frequency band (Herrmann, Munk, & Engel, 2004; Schneider et al., 2008). It was extracted by using the wavelet transform at the target frequency (f) (Lachaux et al., 2000). The frequency ranged from 2 Hz to 45 Hz in 1 Hz steps. To avoid problems due to the sample size bias, for each infant, the number of epochs was made the same for the match and mismatch conditions by randomly selecting the same number of epochs. EEG signal $s(t)$ was convolved with the complex Morlet's wavelet defined by:

$$w(t,f) = \sqrt{f}\exp(-t^2/2\sigma_t^2)\exp(i2\pi ft),$$

as a function of time (t) and frequency (f). The Morlet wavelet is characterized solely by σ_t, which sets the number of cycles of the wavelet: $nco = 6f\sigma_t$. We chose nco to be 8 (Lachaux et al., 2000).

To detect auditory event-related changes in amplitude, we first computed the instantaneous amplitude of EEG signal from electrode n by deriving the length of the convolved signal as follows:

$$A_n(t) = |w(t,f) * s_n(t)|.$$

Next, we averaged the instantaneous amplitude $A_n(t)$ across all trials and obtained averaged amplitude $AMP_n(t)$.

Finally, we standardized the averaged amplitude relative to the pre-stimulus baseline period (600 msec−100 msec before the visual onset) for each electrode and frequency. Standardized amplitude values for each time point t [$AMPz(t)$], were computed as follows:

$$AMP_z(t) = \frac{AMP(t) - AMP_{Bmean}}{AMP_{Bsd}}$$

where AMP_{Bmean} and AMP_{Bsd} are, respectively, the mean and standard deviation of the AMPs computed from the baseline period at each frequency. The resulting index, $AMPz$, indicates standardized changes in the direction of increased amplitude (positive values) or decreased amplitude (negative values).

2.5.2. Phase synchronization analyses

To estimate large-scale synchronization in brain activity, we quantify phase synchronization of neural oscillations by calculating phase-locking values (PLV), which reflect the consistency of phase differences between two EEG electrodes (Lachaux et al., 2000; Rodriguez et al., 1999).

The instantaneous phase of EEG signals was extracted by using the same wavelet transform procedure as in 2.5.1, with which EEG signal $s(t)$ was convolved. We computed the instantaneous phase ϕ_n of EEG signal from electrode n by deriving the argument of the convolved signal:

$$\exp(i\phi(t,f)) = w(t,f) * s_n(t)/|w(t,f) * s_n(t)|.$$

Finally, we computed the PLV to estimate the degree of phase synchronization between EEG phase signals as,

$$PLV(t) = \frac{1}{M}\left|\sum_{m=1}^{M}\exp(i\theta(t,m))\right|,$$

where $\theta(t,m) = \phi_1(t,m) - \phi_2(t,m)$, ϕ_1 and ϕ_2 are the instantaneous phases of EEG time series from electrodes 1 and 2 at time t for the m-th trial (Lachaux et al., 2000; Rodriguez et al., 1999). M is the total number of epochs included in the calculation. The resulting PLV takes a value between 0 (random phase difference, no phase synchronization) and 1 (constant phase difference, perfect phase synchronization).

To detect auditory event-related changes in synchrony, we standardized the PLV relative to the pre-stimulus baseline period (600 msec−100 msec before the visual onset) for each electrode pair and frequency. Standardized PLV values for each time point t, $PLVz(t)$ (Rodriguez et al., 1999), were computed as follows:

$$PLV_z(t) = \frac{PLV(t) - PLV_{Bmean}}{PLV_{Bsd}}$$

where PLV_{Bmean} and PLV_{Bsd} are, respectively, the mean and standard deviation of the PLVs computed from the baseline period at each frequency. The resulting index, $PLVz$, indicates standardized changes in the direction of increased synchronization (positive values) or decreased synchronization (negative values).

2.5.3. ERP analysis

The EEG signal was re-filtered off-line with a zero phase shift digital band-pass filter ranging from .3 to 30 Hz, and re-referenced to the average of left and right mastoid channels (A1, A2). Artifact rejection was performed automatically by rejecting trials with a potential exceeding ±200 μV. There was a minimum of 21 valid epochs per condition in every infant participant (mean: 47.6 epochs in the match condition and 46.7 epochs in the mismatch condition). Epochs ranged from −950 to 1000 msec after the auditory onset and baseline correction was applied in the interval −950 to −550 msec (i.e., from 400 to 0 msec before the onset of the visual stimulus).

We calculated mean amplitudes within a time window of 350–550 msec after the auditory onset over the central regions of the scalp (i.e., C3, Cz, and C4) to evaluate the N400 effect. A two-way analysis of variance (ANOVA) (two sound-symbolic matching conditions × three electrodes) on the mean amplitudes in the time-window was conducted.

3. Results

3.1. Amplitude change

We computed $AMPz$ on an individual basis. The statistical group analyses were performed on $AMPz$ time-frequency diagrams. Those diagrams were grouped across infants by

condition and analysed by a permutation test in search of time-frequency windows showing significant differences between the two conditions (Burgess & Gruzelier, 1999; Melloni et al., 2007). Specifically, the significance level of group *AMPz* difference (real difference) was tested in a pseudo-random distribution of group differences obtained by randomly shuffling ($N = 10{,}000$) the label of conditions (i.e., match or mismatch) of time-frequency diagrams within each infant. The statistical effects of multiple comparisons were controlled by FDR (False Discovery Rate; see Benjamini & Hochberg, 1995) by the number of electrodes (i.e., 9 electrodes). We considered a measured *AMPz* difference above the (FDR-corrected) 97.5th percentile or below the 2.5 percentile of the pseudo-random distribution of *AMPz* differences to be significant.

Fig. 3(a) displays the resulting standardized *AMP* (*AMPz*) averaged across all 9 electrodes and all infants for the match and mismatch conditions, and the differences in *AMPz* between the two conditions. Fig. 3(b) presents a topographic map showing significant *AMPz* differences between the two conditions lasting more than .86 frequency cycles in each time window. The .86 frequency cycle criterion was chosen in such a way that the type I error does not occur in the baseline time window, where no difference between the match and mismatch conditions should be observed.

The results revealed an increase of gamma-band (34–37 Hz) amplitude in the match condition as compared to the mismatch condition in the 1–300 msec time window, which is earlier than the typical N400 time window (e.g., around 400 msec). The increased gamma-band activity for the sound-symbolically matched shape–sound pairs in the early time window is consistent with previous EEG amplitude studies on multi-sensory integration in adults (e.g., Schneider et al., 2008; in Schneider et al., gamma-band activity increased for matched audio-visual stimuli at around 100–200 msec and 40–50 Hz), and also with results reported by Csibra et al. (2000), in which an increased gamma-band activity (at around 40 Hz) was observed for visual feature binding in 8-month-old infants at 180–320 msec after stimulus onset. The gamma-band increase was observed at the centro-parietal regions (electrodes C4, P3, Pz, and P4). This is also similar to the study of Schneider et al. (2008), in which gamma-band increase was observed at medial central regions.

The early increase of gamma-band EEG amplitude for sound-symbolically matched sound-shape pairs was subsequently followed by beta- (and theta-) band increases in the 301–600 time window and by gamma- (and theta-) band increases in the 601–900 msec time window both for sound-symbolically mismatched sound-shape pairs. Beta-band activity, which is sometimes accompanied by amplitude increase in the theta, alpha and gamma band, is known to be involved in perceptual cross-modal processing (Senkowski et al., 2008, for a review). Gamma-to-beta shift in neural activities is reported in some studies on both uni-modal and cross-modal perceptual integration, and some researchers argue that gamma- and beta-band activity correspond to bottom-up perceptual processing and stimulus-driven salience signalling, respectively (Kisley & Cornwell, 2006). The increased beta-band activity for sound-symbolically mismatched sound-shape pairs as compared to sound-symbolically matched pairs may indicate that infants attended to the stimulus pairs more closely when they were sound-symbolically mismatched than matched.

3.2. *Phase synchronization*

We computed *PLVz* on an individual basis. The statistical group analyses were performed on *PLVz* time-frequency diagrams by using the same permutation test procedure as for the amplitude change (*AMPz*) analyses, except that the FDR control of multiple comparisons of statistical effects was made by the number of electrode pairs (i.e., 36 pairs) this time.

Fig. 3(d) displays the resulting standardized *PLV* (*PLVz*) averaged across all 36 electrode pairs and all infants for the match and mismatch conditions. Prominent large-scale synchronization was observed immediately after the auditory onset (0 msec) across the alpha-beta bands (9–15 Hz) in both conditions. In the match condition, however, active phase synchronization was no longer evident from about 300 msec after the auditory onset. In the mismatch condition, in contrast, phase synchrony was stronger and more durable in the later time windows (300 msec onwards) than in the match condition. When comparing the two conditions, a marked difference in large-scale phase synchronization was found in the beta band (12–15 Hz), which is in accordance with previous findings reporting the involvement of beta-band amplitude increase and coherence in multi-sensory integration (Senkowski et al., 2008).

Fig. 3(c) presents a topographical map showing significant PLV_z difference between the two conditions lasting more than .96 frequency cycles in each time window. The .96 frequency cycle criterion was chosen in such a way that a type I error was not found in the baseline time window, where no difference between the match and mismatch conditions should be observed. A statistically significant difference was found between the match and mismatch conditions in the latter two time windows (301–600 msec, 601–900 msec). In these time windows, phase synchronization increased for sound-symbolically mismatched sound-shape pairs than for sound-symbolically matched pairs in the beta band (14–15 Hz), most prominently between electrode P3 (and C3) and other electrodes over the left scalp. The N400 time-window coincided with the time period in which the most prominent difference in synchronization between matching and mismatching conditions was found. See Supplementary Fig. S1 for a topographical map showing significant *PLVz* for the match and mismatch conditions as compared to pre-stimulus baseline.

Spurious phase synchrony of EEG signals could arise from volume conduction due to a single dipole activity. To rule out this possibility, we also analysed the distribution of instantaneous phase differences between pairs of electrodes. If spurious synchrony had been caused by volume conduction, distributions narrowly centred on zero and pi (Melloni et al., 2007) would have been observed. However, the results indicated that this was not the case, as scattered distributions were observed.

3.3. *ERP*

As Fig. 4 shows, we identified a typical adult-like N400 response in infants. ERPs to sound-symbolically mismatched stimuli

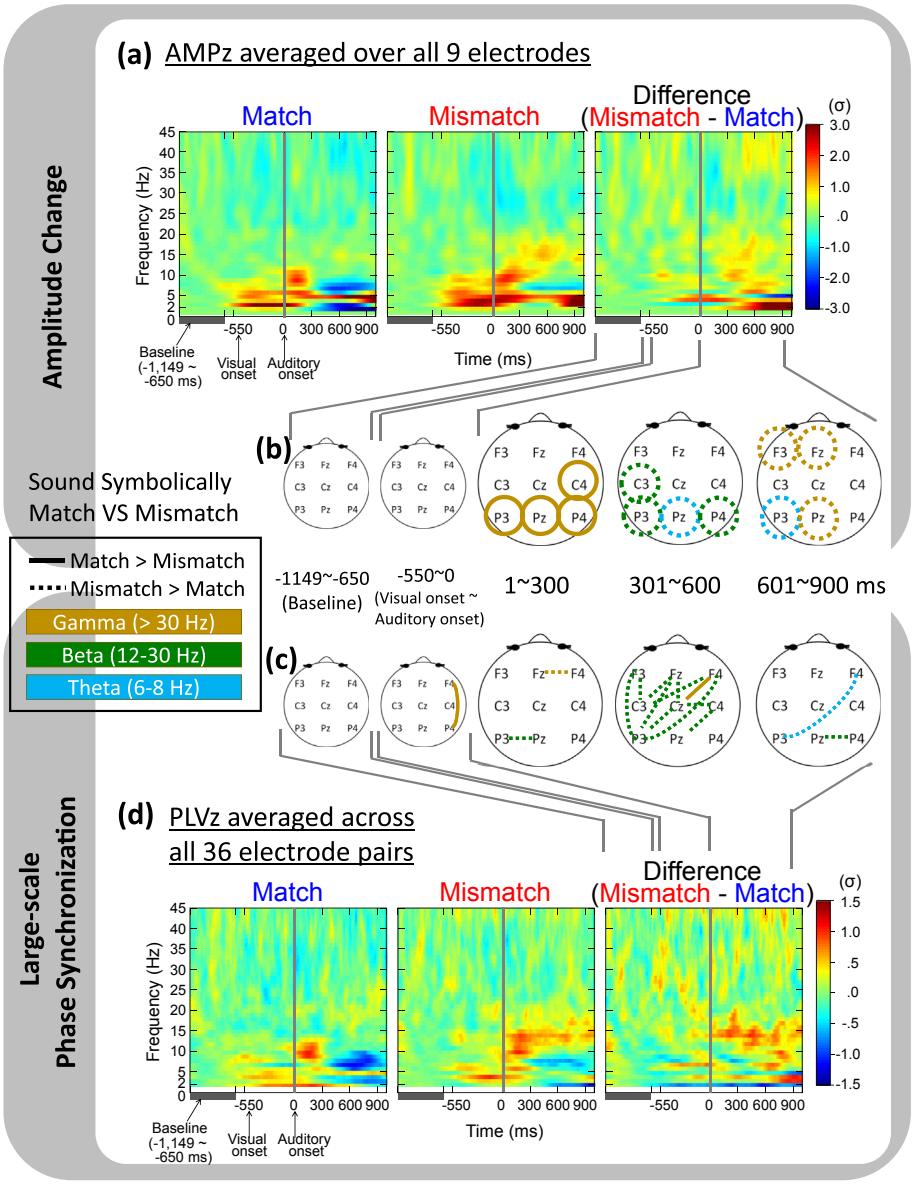

Fig. 3 – (a) Time-frequency diagrams of grand average *AMPz* (in the standard deviation unit). *AMPz* was averaged across all 9 electrodes and 19 infants for the match (a, left) and mismatch (a, middle) conditions. Positive *AMPz* (in warm colours) indicates standardized changes in the direction of increased synchronization and negative *AMPz* (in cold colours) indicates decreased synchronization. (a, right) Difference in grand average *PLVz* across the match and mismatch conditions. Note that soon after the auditory onset (<300 msec), the grand average *AMPz* for the match condition increased than those for the mismatch condition in the gamma frequency range (at around 35 Hz). (b) Topographic maps of significant difference in *AMPz* between the match and mismatch conditions. The solid line circles and the dotted line circles show significantly ($p < .05$, FDR corrected, $N = 19$) higher *AMPz* in the match than in the mismatch condition and in the mismatch than in the match condition, respectively. (c) Topographic maps of significant difference in *PLVz* between the match and mismatch conditions. The solid lines and the dotted line show significantly ($p < .05$, FDR corrected, $N = 19$) higher *PLVz* in the match than in the mismatch condition and in the mismatch than in the match condition, respectively. (d) Time-frequency diagrams of grand average *PLVz* (in the standard deviation unit). *PLVz* was averaged across all 36 electrode pairs and 19 infants for the match (d, left) and mismatch (d, middle) conditions. Positive *PLVz* (in warm colours) indicates standardized changes in the direction of increased synchronization and negative *PLVz* (in cold colours) indicates decreased synchronization. (d, right) Difference in grand average *PLVz* across the match and mismatch conditions. Note that after the auditory onset, the grand average *PLVz* for the mismatch condition display more sustained pattern than for the match condition in the alpha-to-beta frequency range.

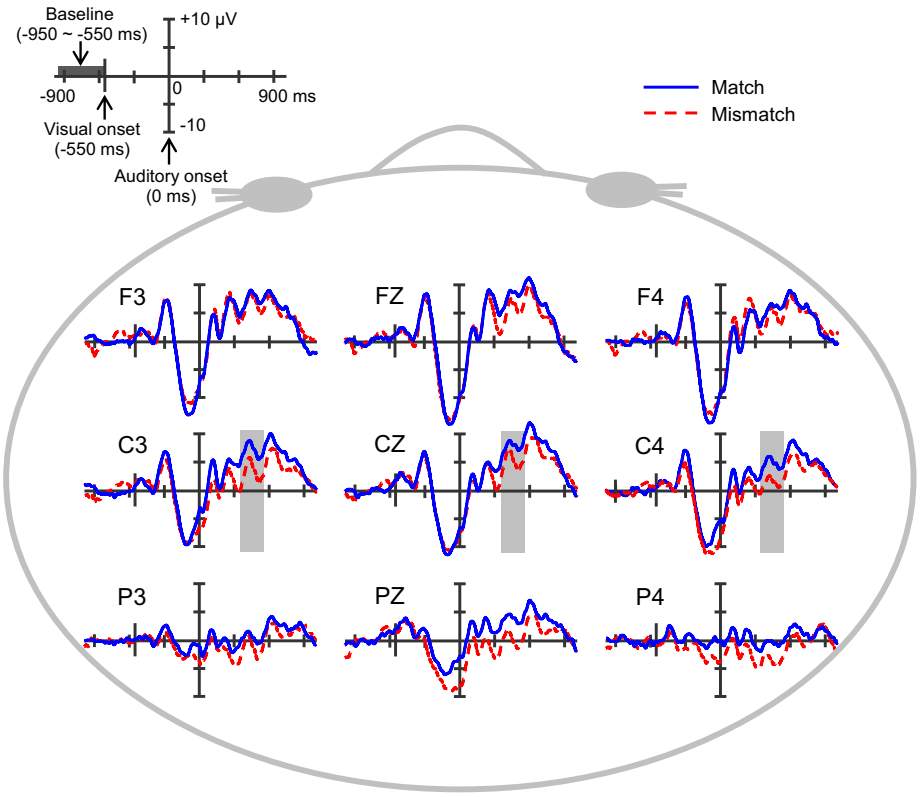

Fig. 4 – ERPs elicited by sound-symbolically matched (blue solid lines) and mismatched (red dotted lines) stimuli ($N = 19$). ERPs to sound-symbolically mismatched stimuli were significantly more negative going than those to sound-symbolically matched stimuli at around 350–550 msec after the auditory onset over the central regions of the scalp, i.e., C3, Cz, and C4 (shaded areas), which correspond to the typical time-window and sites for the N400 effect.

were more negative going than those to sound-symbolically matched stimuli at around 350–550 msec after the auditory onset over the central regions of the scalp, i.e., C3, Cz, and C4, which correspond to the typical time-window and sites for the N400 effect (Kutas & Federmeier, 2011). A two-way ANOVA (two sound-symbolic matching conditions × three electrodes) on the mean amplitudes in the time window revealed a main effect of sound-symbolic matching [$F(1,18) = 8.47$, $p < .01$, two-tailed, $\eta^2 = .03$, $N = 19$; all data were normally distributed (all $Ds < .16$ and $ps > .62$, Kolmogorov–Smirnov test)]. No statistical differences between the two conditions were found in other time windows including earlier time windows (e.g., 1–300 msec, in which the differences between conditions were found in the amplitude change analysis) over any scalp regions [frontal (i.e., F3, Fz, and F4), central (i.e., C3, Cz, and C4), and parietal (i.e., P3, Pz, and P4)].

4. Discussion

This study investigated the neural mechanism for processing novel word–shape pairs with or without sound symbolism in 11-month-old infants. There were three key findings:

First, amplitude change assessed by *AMP* increased for sound-symbolically matched sound-shape pairs more than for sound-symbolically mismatched pairs in the gamma band and in an early time window (1–300 msec), consistent with previous infant studies showing that perceptual processing modulates oscillation amplitude in the gamma band in the same time window (Csibra et al., 2000). Thus, the results from the amplitude change analysis suggest that sound symbolism is processed as a perceptual binding in 11-month-old infants.

Second, phase synchronization of neural oscillations assessed by *PLV* increased, as compared to the baseline period, significantly more in the mismatch condition than in the match condition. This effect was observed in the beta-band and most pronounced over left-hemisphere electrodes during the time window (301–600 msec) in which the N400 effect was detected in ERP. The time course of large-scale synchronization suggests that cross-modal binding was achieved quickly in the match condition, but sustained effort was required in the mismatch condition and seemed to involve left-lateralized structures. The stronger inter-regional communication in the left hemisphere is compatible with the idea that the language-processing network in the left hemisphere (Mesulam, 1990; Springer et al., 1999) is recruited for processing the sound-shape pairings. For the first time in infants, the current study shows modulation of large-scale synchronization by cognitive processing in the time window coinciding with that in which semantic (or higher cognitive level) processing has been identified in adults.

When adult participants are presented with real words and non-words in isolation, real words elicit stronger EEG coherence in the beta-band in comparison to the resting state, but

non-words do not (von Stein, Rappelsberger, Sarnthein, & Petshe, 1999). This indicates that lexical processing induces beta-band synchronization in adults. The beta-band increase in phase synchronization found in our infants suggests that the same neural network may be recruited for processing words already at the age of 11 months.

The third key finding is that the N400 component was significantly larger for sound symbolically mismatching than matching pairs. The difference in ERP amplitude between the match and mismatch condition suggests that 11-month-olds' brain sensitively responds to congruency of sound-shape correspondences. Furthermore, the timing and topography of this ERP modulation is strikingly similar to the typical N400 effect (Kutas & Federmeier, 2011). Although there is widespread agreement in the literature that the N400 response reflects semantic integration difficulty both in adults and infants (Friedrich & Friederici, 2005, 2011; Kutas & Federmeier, 2011; Parise & Csibra, 2012), the neural mechanism underlying N400 is not perfectly understood (Kutas & Federmeier, 2011), especially in infants. In our case, however, the results from the amplitude change in the earlier time window along with the large-scale posterior-anterior synchrony observed in the beta band over the left hemisphere in the N400 time window jointly suggest that N400 modulation reflects the detection of an anomaly at a conceptual rather than perceptual level. Indeed, when visual shape and spoken word were sound-symbolically mismatched, it was more difficult for infants to integrate the two and establish the pairing. In other words, sound symbolism may help infants to acquire the concept of word from novel sound-referent pairing.

This study goes beyond effects of sound symbolism previously demonstrated in infant behavioural measures (Maurer et al., 2006; Ozturk et al., 2013; Peña et al., 2011; Walker et al., 2010), as it revealed the neural processes linking perceptual cross-modal processing and language development. The amplitude change, phase synchronization, and ERP results jointly indicate that, while it is processed in a cross-modal perceptual network, sound symbolism triggers semantic processing in the left hemisphere mapping speech sounds to visually presented referents.

Sound symbolism may serve as an important bootstrapping mechanism for establishing referential insights for speech sounds. Although infants may start to associate words and referents as early as 6 months of age (Bergelson & Swingley, 2012), this process is initially slow and effortful. Reports in the literature show that, up to 13 months, infants are not so efficient in picking up the referent in preferential looking, but that at 14 months, they become qualitatively different and perform much better (Bergelson & Swingley, 2012; Werker et al., 1998). In other words, young infants need much more scaffolding to establish word referent associations than older infants. Sound symbolism may be a helping cue derived from a naturally endowed biological capacity to map speech sounds to perceptual properties (Gogate & Hollich, 2010). After a phase in which sound symbolism helps infants to become aware of the meaningful association between speech sounds and referents, infants may intentionally seek to associate speech sounds to referents, which would in turn lead to the realization that not all sound–referent pairs have a close sound–symbolic relationship. This process is likely to prompt referential insight before the establishment of arbitrary word–meaning relationships.

As this study is the first to explore the neural processing of sound symbolism in the infants' brain, several limitations should be acknowledged. First, the generalizability of our results will need to be examined using large sets of word-referent pairs including in other perceptual domains than vision. Second, although the large-scale synchronization in the beta band found in the infants in this study is consistent with the pattern found in previous study in adults (von Stein et al., 1999), developmental trajectory of beta-band synchrony needs much more investigation. Only a few studies have investigated how large-scale neural synchronization networks, mediating inter-regional communication and brain functions, develop and mature in humans. Uhlhaas and colleagues (Uhlhaas, Roux, Rodriguez, Rotarska-Jagiela, & Singer, 2010; Uhlhaas et al., 2009) analysed the development of functional networks by measuring EEG oscillations and synchrony during a face perception task in participants ranging in age from 6 to 21 years. Their results suggest that developmental improvements in cognitive performance are accompanied by increases in gamma-band power and beta-band neural synchrony. Although they did not test younger children, their results underscore the importance of development in large-scale beta-band synchrony in cognitive processing. Further investigation is necessary to understand how the pattern of whole brain communication develops in the course of language development and how they map to cognitive functions in language processing.

Despite its limitations, this study methodologically expands the horizon of developmental neuroscience research. Studies addressing the neural processing of semantic information in the infant brain are still sparse. There are a handful of studies demonstrating the N400 effect under certain conditions in infants (Friedrich & Friederici, 2005, 2011; Parise & Csibra, 2012; Thierry, 2005; Torkildsen et al., 2008). However, no study has examined connectivity amongst different regions in the infant brain when language processing takes place. This study is the first step toward understanding how the infant brain creates networks when establishing word-referent associations.

Acknowledgements

This research was supported by MEXT KAKENHI (#15300088, #22243043, Grant-in-Aid for Scientific Research on Innovative Areas #23120003) to M.I. and H.O., MEXT KAKENHI (#21120005) and JST PRESTO to K.K., MEXT GCOE program to Tamagawa University, BBSRC Research Development Fellowship (BB/G023069/1) to S.K., Economic and Social Research Council (ES/E024556/1) and European Research Council (ERC-SG-209704) to G.T, and Grant-in-Aid for JSPS Research Fellows (#23-2872) to M.A. We thank Yumi Nakagawa,Yuji Mizuno, Junko Kanero and Mamiko Arata for help in data collection and analysis, and Marilyn Vihman for comments on an earlier version of the manuscript. M.A. and M.I. are joint first authors. G.T. and S.K. made equal contributions. The authors declare no competing financial interests.

Supplementary data

Supplementary data related to this article can be found at http://dx.doi.org/10.1016/j.cortex.2014.08.025.

REFERENCES

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological), 57*, 289–300.

Bergelson, E., & Swingley, D. (2012). At 6 to 9 months, human infants know the meanings of many common nouns. *Proceedings of the National Academy of Sciences of the USA, 109*, 3253–3258.

Burgess, A., & Gruzelier, J. (1999). Methodological advances in the analysis of event-related desynchronization data: reliability and robust analysis. In G. Pfurtscheller, & F. H. Lopes da Silva (Eds.), *Handbook of electroencephalography and clinical neurophysiology, revised series* (pp. 139–158). Amsterdam: Elsevier.

Csibra, G., Davis, G., Spratling, M. W., & Johnson, M. H. (2000). Gamma oscillations and object processing in the infant brain. *Science, 290*, 1582–1585.

Engel, A., & Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends in Cognitive Sciences, 5*, 16–25.

Fennell, C. T., & Werker, J. F. (2003). Early word learners' ability to access phonetic detail in well-known words. *Language and Speech, 46*, 245–264.

Friedrich, M., & Friederici, A. D. (2005). Lexical priming and semantic integration reflected in the event-related potential of 14-month-olds. *NeuroReport, 16*, 653–656.

Friedrich, M., & Friederici, A. D. (2011). Word learning in 6-month-olds: fast encoding – weak retention. *Journal of Cognitive Neuroscience, 23*, 3228–3240.

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends in Cognitive Sciences, 9*, 474–480.

Gogate, L. J., & Hollich, G. (2010). Invariance detection within an interactive system: a perceptual gateway to language development. *Psychological Review, 117*, 496–516.

Herrmann, C. S., Munk, M. H. J., & Engel, A. K. (2004). Cognitive functions of gamma-band activity: memory match and utilization. *Trends in Cognitive Sciences, 8*, 347–355.

Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B, 369*. http://dx.doi.org/10.1098/rstb.2013.0298.

Imai, M., Kita, S., Nagumo, M., & Okada, H. (2008). Sound symbolism facilitates early verb learning. *Cognition, 109*, 54–65.

Imai, M., Miyazaki, M., Yeung, H. H., Hidaka, S., Kantartzis, K., Okada, H., et al. *Sound symbolism facilitates word learning in 14-month-olds* (under review).

Kantartzis, K., Imai, M., & Kita, S. (2011). Japanese sound-symbolism facilitates word learning in English-speaking children. *Cognitive Science, 35*, 575–586.

Kawasaki, M., Kitajo, K., & Yamaguchi, Y. (2010). Dynamic links between theta executive functions and alpha storage buffers in auditory and visual working memory. *European Journal of Neuroscience, 31*, 1683–1689.

Kisley, M. A., & Cornwell, Z. M. (2006). Gamma and beta neural activity evoked during a sensory gating paradigm: effects of auditory, somatosensory and cross-modal stimulation. *Clinical Neurophysiology, 117*, 2549–2563.

Kitajo, K., Doesburg, S. M., Yamanaka, K., Nozaki, D., Ward, L. M., & Yamamoto, Y. (2007). Noise-induced large-scale phase synchronization of human-brain activity associated with behavioural stochastic resonance. *Europhysics Letters, 80*(40009), 1–6.

Köhler, W. (1947). *Gestalt psychology* (2nd ed.). New York: Liveright.

Kutas, M., & Federmeier, K. D. (2011). Thirty years and counting: finding meaning in the N400 component of the event-related brain potential (ERP). *Annual Review of Psychology, 62*, 621–647.

Lachaux, J.-P., Rodriguez, E., Le Van Quyen, M., Lutz, A., Martinerie, J., & Varela, F. J. (2000). Studying single-trials of phase synchronous activity in the brain. *International Journal of Bifurcation and Chaos, 10*, 2429–2439.

Maurer, D., Pathman, T., & Mondloch, C. J. (2006). The shape of boubas: sound–shape correspondences in toddlers and adults. *Developmental Science, 9*, 316–322.

Melloni, L., Molina, C., Pena, M., Torres, D., Singer, W., & Rodriguez, E. (2007). Synchronization of neural activity across cortical areas correlates with conscious perception. *The Journal of Neuroscience, 27*, 2858–2865.

Mesulam, M. (1990). Large-scale neurocognitive networks and distributed processing for attention, language, and memory. *Annals of Neurology, 28*, 597–613.

Miyazaki, M., Hidaka, S., Imai, M., Yeung, H.,H., Kantartzis, K., Okada, H., et al. (2013). The facilitatory role of sound symbolism in infant word learning. In *Proceedings of the 35th Annual Meeting of the Cognitive Science Society*.

Ozturk, O., Krehm, M., & Vouloumanos, A. (2013). Sound symbolism in infancy? Evidence for sound-shape cross-modal correspondences in 4-month-olds. *Journal of Experimental Child Psychology, 114*, 173–186.

Parise, E., & Csibra, G. (2012). Electrophysiological evidence for the understanding of maternal speech by 9-month-old infants. *Psychological Science, 23*, 728–733.

Peña, M., Mehler, J., & Nespor, M. (2011). The role of audiovisual processing in early conceptual development. *Psychological Science, 22*, 1419–1421.

Ramachandran, V. S., & Hubbard, E. M. (2001). Synaesthesia - a window into perception, thought and language. *Journal of Consciousness Studies, 8*, 3–34.

Rodriguez, E., George, N., Lachaux, J.-P., Martinerie, J., Renault, B., & Varela, F. J. (1999). Perception's shadow: long-distance synchronization of human brain activity. *Nature, 397*, 430–433.

Schneider, T. R., Debener, S., Oostenveld, R., & Engel, A. K. (2008). Enhanced EEG gamma-band activity reflects multisensory semantic matching in visual-to-auditory object priming. *NeuroImage, 42*, 1244–1254.

Senkowski, D., Schneider, T. R., Foxe, J. J., & Engel, A. K. (2008). Crossmodal binding through neural coherence: implications for multisensory processing. *Trends in Neurosciences, 31*, 401–409.

Springer, J. A., Binder, J. R., Hammeke, T. A., Swanson, S. J., Frost, J. A., Bellgowan, P. S., et al. (1999). Language dominance in neurologically normal and epilepsy subjects. A functional MRI study. *Brain, 122*, 2033–2045.

Stager, C. L., & Werker, J. F. (1997). Infants listen for more phonetic detail in speech perception than in word-learning tasks. *Nature, 388*, 381–382.

von Stein, A., Rappelsberger, P., Sarnthein, J., & Petsche, H. (1999). Synchronization between temporal and parietal cortex during multimodal object processing in man. *Cerebral Cortex, 9*, 137–150.

Tallon-Baudry, C., Bertrand, O., Delpuech, C., & Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 Hz visual responses in human. *The Journal of Neuroscience, 16*, 4240–4249.

Thierry, G. (2005). The use of event-related potentials in the study of early cognitive development. *Infant and Child Development, 14*, 85–94.

Torkildsen, J. V. K., Svangstu, J. M., Friis Hansen, H., Smith, L., Simonsen, H. G., Moen, I., et al. (2008). Productive vocabulary size predicts ERP correlates of fast mapping in 20-month-olds. *Journal of Cognitive Neuroscience, 20*, 1266–1282.

Uhlhaas, P. J., Roux, F., Rodriguez, E., Rotarska-Jagiela, A., & Singer, W. (2010). Neural synchrony and the development of cortical networks. *Trends in Cognitive Sciences, 14*, 72–80.

Uhlhaas, P. J., Roux, F., Singer, W., Haenschel, C., Sireteanu, R., & Rodriguez, E. (2009). The development of neural synchrony reflects late maturation and restructuring functional networks in humans. *Proceedings of the National Academy of Sciences of the USA, 106*, 9866–9871.

Varela, F., Lachaux, J.-P., Rodriguez, E., & Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nature Reviews Neuroscience, 2*, 229–239.

Walker, P., Bremner, J. G., Mason, U., Spring, J., Mattock, K., Slater, A., et al. (2010). Preverbal infants' sensitivity to synaesthetic cross-modality correspondences. *Psychological Science, 21*, 21–25.

Ward, L. M. (2003). Synchronous neural oscillations and cognitive processes. *Trends in Cognitive Sciences, 7*, 553–559.

Werker, J. F., Cohen, L. B., Lloyd, V. L., Casasola, M., & Stager, C. L. (1998). Acquisition of word–object associations by 14-month-old infants. *Developmental Psychology, 34*, 1289–1309.